Impact Factor: 9.9 ISSN-L: 2544-980X

Data Analysis: Types, Process, Methods, Techniques and Tools

Khakimov Akhror Abdimakhamadovich [1]

Abstract: Academic theses at the bachelor's and master's levels in nursing, physiotherapy, occupational therapy, public health and related fields often take the form of general reviews of published research on a relevant clinical issue. While there are many guides to doing literature searches and evaluating article quality, there are to our knowledge no published detailed guides of how to do the actual data analysis in such general literature reviews, particularly that are applicable to students. This article seeks to describe a systematic method of data analysis appropriate for undergraduate research theses, where the data consists of the results from available published research. We present a step-by-step guide with authentic examples and practical tips.

Keywords: academic theses, data analysis, general literature reviews, healthcare education, practical guide.

The central task for the author of any general literature review is to analyse the results of multiple scientific studies in order to describe the state of knowledge about a particular topic, in order to draw conclusions with clinical applications. This article seeks to describe a systematic method of data analysis appropriate for undergraduate theses, where the data consists of the results from available published research. Although the methods described here are usable by any researchers conducting a general literature review, the article is addressed to students (and their supervisors) in the health sciences. A general literature review starts with formulating a research question, defining the population, and conducting a systematic search in scientific databases, steps that are well-described elsewhere.1,2,3 Once students feel confident that they have thoroughly combed through relevant databases and found the most relevant research on the topic, however, what is arguably the hardest part of the process remains. How should one go about synthesizing, analysing and presenting the results of different studies on a topic? This is an especially challenging task when these studies may not themselves ask the exact same research question you the student are asking, even if they contain findings that answer your research question. That the studies may use vastly varying methodologies, from interview studies to randomized trials, also poses a challenge to the task of synthesising and analysing the data. We have discovered that students lack a structured guide to the actual analysis that a general literature review requires. Below we present a step-by-step guide for analysing data for two different types of research questions. The data analysis methods described here are based on basic content analysis as described by Elo and Kyng€as4 and Graneheim and Lundman,5 and the integrative review as described by Whittemore and Knafl,6 but modified to be applicable to analysing the results of published studies, rather than empirical data. The methods described here are inductive, that is, they do not describe how to use a pre-existing model or theory to analyse data, but instead describe how to find patterns and answers in the data without preconceived theories or frameworks for analysis.

Once you have completed searches of the databases and identified all studies that answer your research question and fit within the inclusion and exclusion criteria you have chosen (for example only studies on people aged over 65 years, or no studies of people with disabilities), it is time to get started on the data analysis. The first step is to get an overview of all identified studies by making an article matrix where you describe each study's aim, methods and results, especially those results that are relevant to your particular research question. Table 1 shows examples of how this might look for an experimental, randomized controlled trial study; for an interview study; and for a study based on a questionnaire. All of these studies will be used in examples later in this article; the first article for a study about inserting

[1] Ferghana branch of the Tashkent University of Information Technologies named after Muhammad al-Khorezmi

peripheral vein catheters (PVCs), and the second two for a study about patient experiences of isolation wards. The next steps in the data analysis process differ depending upon the type of research question you have and the designs of the studies you will be analysing. When your research question (aim) is about effect, risk, association or prevalence Research questions about whether or not a certain treatment, intervention or approach is effective are answered primarily by experimental studies, including randomized controlled trials (RCT) and quasi-experimental studies. Research questions about risk factors, associations and prevalence are answered primarily by cross-sectional and longitudinal studies. The data analysis process for these kinds of quantitative studies involves three steps: 1. Identifying data that answers your research question (aim), in this case largely numerical data that must be 'dug' out of each study's results. 2. Organizing the data in a thematic way. 3. Synthesizing, analysing, and presenting the data. This section will now go through each of these steps, using the example of the following research question: What strategies are effective for inserting peripheral venous catheters in patients with difficult intravenous access? In the development of this research question you have hopefully already determined whether you only want to examine strategies that lead to success in insertion of a PVC, or whether you also want to include findings about patient experiences of pain during the procedure, how the skill of the provider affects insertion success, or other aspects of what 'success' means for a PVC insertion With your research question and its exact terms in mind, you begin the process of data analysis by clearly identifying the actual results of each study that answer your research question. This entails separating the wheat from the chaff; that is, picking out the study results that are relevant to your research question, and ignoring results that are not. Practical tip: Tape your research question to your computer so that you don't lose sight of it. You can also write it on the top margin of your draft during the writing period so that it is visible on every page (just don't forget to remove it at the end). Be aware that it is not always easy to identify the actual results of any study that answer your particular research question. For example, if you are interested in effective strategies for inserting PVCs in adult patients with difficult vascular access then you need to ignore study results that a) concern children and b) concern other kinds of vascular access or catheterization. Box 1 shows an example of text from the results section of an article with the information that answers your research question about effective strategies for inserting PVCs in people with difficult intravenous access. The parts of the text that answer your research question are underlined. Note that figuring out which data is relevant can be difficult. The relevant data will be presented in the Results section, may be mentioned in the Abstract, and is often summarized at the very beginning of the Discussion section of an article. (If you find results conveniently summarized at the beginning of a Discussion section, just make sure to avoid including the actual 'discussion' of these results by the authors, where they compare their results to those of other studies or give their own interpretations and opinions about their results.) The Abstract seldom describes findings in enough detail; sometimes the very results that are key to answering your particular research question are not even mentioned in the Abstract. Sometimes the findings are explained clearly in the text in the Results section of a study, but sometimes one may need to look at tables, graphs and figures to find the relevant information. Quantitative studies tend to contain a lot of statistical data, and not all of it is relevant in a general literature review, especially at the undergraduate level. Tables that show the baseline characteristics of the population studied, for example, are important for providing the study with validity (roughly, accuracy and trustworthiness), but may not be relevant to your synthesis of the study's findings. In general, authors focus on results that are statistically significant described in a study's Discussion section, but you can also look for aspects of a study that the authors themselves do not specifically mention, but that you notice may have had an effect on their results. A weakness could be, for example, that a high number of participants had multiple health problems affecting the results, or that a company that sold ultrasound equipment financed the study, or other aspects that might detract from (or add to) the validity and reliability (roughly, accuracy of measurement) of the results. Table 2 shows an example of how data from our PVC study could be organized into a table to create a good overview of the studies' results as well as the studies' strengths and weaknesses.

The challenge of synthesising and analysing quantitative data is that studies of the same phenomenon tend to study the phenomenon in different ways and to measure their results in different ways. For

example, to measure outcomes, two studies measured the percentage of successful attempts,7,11 whereas another study divided patients into groups with different degrees of difficulty of access.10 Yet another study, a cohort study, lacking a control group, found merely that 93% of patients, deemed to have difficult access, were able to have an IV line inserted when ultrasound was used.12 We do not know what the success rate would have been for a matched control group where ultrasound was not used; only that previous studies have rarely found such a high success-of-insertion rate for patients with difficult intravenous access when ultrasound is not used.

The same process is followed for each of the studied procedures: heat, alternative therapies, etc. That is, the overall results of the studies of heat's effects on easing catheter insertion, for example, are presented, and then the specifics of each study that examined heat's effects are presented. Depending on the wording of your research aim, you might want to include other types of results, such as how much pain patients felt when different procedures were used, or whether the use of ultrasound was more successful with experienced and trained providers, which some studies examined. It is also important to note each study's validity, that is, the validity of the study's results. Not all results are equal, because some studies are better than others. For example, a study with a large sample, careful randomization, and few confounding factors, will produce a more trustworthy result than a study with a smaller sample, selection bias, no randomization, and many confounding factors. The validity and reliability of particular studies can be woven into your synthesis of results, as in the example in Box 2, or discussed separately at the end of your results section